Nucleic Acids Research, Vol. 19, No. 15 4147–4151

Characterization of human 5S rRNA genes

Pernille Dissing Sørensen and Sune Frederiksen*
Department of Biochemistry B, Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

Received May 1, 1991; Revised and Accepted July 4, 1991

ABSTRACT

The human 5S rRNA genes are found in clusters of tandem repeated units. We have cloned and partially characterized six restriction fragments from two clusters of 2.3 kb and 1.6 kb repeats, respectively. Four fragments from the cluster of 2.3 kb repeats contain a 5S rRNA gene and one fragment contains a gene variant with an additional nucleotide in the internal control region. A fragment from the 1.6 kb cluster contains a gene and is highly homologous to the 2.3 kb repeats, except for a large deletion in the 3'-flanking region starting 12 bp downstream of the gene. The number of genes and closely related gene variants is found to be 300 – 400 per haploid human genome. 100 – 150 of these are found in 2.3 kb repeats and 5 – 10 are found in 1.6 kb repeats. The total number of 5S rRNA sequences, including pseudogenes, is 1700 – 2000 per haploid genome. The genes and the gene variant are transcribed equally efficient in a HeLa cell extract. If 5'-flanking sequences, including a GC-motif in the – 40 region, are removed from the genes, transcription is reduced with a factor 10 or more, suggesting that sequences upstream of the coding region are important for the level of transcription.

INTRODUCTION

The structure of 5S rRNA genes has been characterized primarily in lower eucaryotes such as Yeast (1), Drosophila (2) and especially Xenopus (3) whereas only limited information is available in higher eucaryotes. With a few exceptions (1,4), the majority of 5S rRNA genes are organized in clusters of tandem repeated units, ranging in size from 375 bp to several kilobases and copy numbers from 160 to 20.000 per haploid genome (2,3,5).

In addition to the genes coding for the 5S rRNA found in ribosomes, most species contain gene variants and pseudogenes differing from the gene by a variable number of substitutions and deletions. Both gene variants and pseudogenes have been found within the clusters or dispersed in the genomes of Xenopus (3), Drosophila (6) and mammalian cells (5,7,8). In general, the pseudogenes are considered transcriptionally inactive, whereas the gene variants, some of which differ from the genes in only one or a few positions, are usually transcribed in vitro. It is not known whether the gene variants are expressed in vivo. So far chicken cells are the only example of a higher eucaryotic cell containing two different 5S rRNA species (9).

A number of 5S rRNA pseudogenes and gene variants have been isolated from rodent cells (5,7,10), but only from Syrian hamster has a bona-fide 5S rRNA gene been cloned (11). The gene was found in a 2.2 Kb BamHI fragment and suggested to be one of 2700 copies of 5S rRNA genes in the hamster genome.

Human cells have been found to contain between 600 and 2000 copies of 5S rRNA genes per haploid genome (12,13). Three pseudogenes and two gene variants have been isolated from the human genome in different restriction fragments (5,8,14).

Recently human 5S rRNA genes have been cloned and sequenced and the genomic structure has been characterized (15,16). In this paper we present additional information on the organization and copy number of human 5S rRNA genes, gene variants and pseudogenes. We also describe the transcription of cloned 5S genes and a gene variant in vitro.

MATERIALS AND METHODS

Cloning

Human placenta DNA was digested with restriction endonucleases SacI or BamHI or SacI + BamHI and then electrophoresed in 0.7% agarose gels. After excision of the 2.3 Kb (SacI), 1.6 Kb (BamHI) and 640 bp (BamHI + SacI) regions, DNA was purified by electroelution, phenol extraction and precipitation from 0.7 M NH_4OAc by ethanol. DNA was ligated into Bluescript M13 plasmids and transformed into E. coli/DH5α made competent with $RbCl_2$ (17). Supercoiled plasmid DNA was purified by the method described by Holmes and Quigly (18) combined with acid-phenol extraction (19) or by the method using alkaline-SDS, LiCl and PEG as described by Sambrook et al. (20). Screening of transformants was carried out with ^{32}P-UTP labelled RNA run-off transcripts synthesized by SP6 transcription of a mouse 5S rRNA gene variant (10).

RNA probes

A mouse 5S rRNA gene variant was isolated in a 729 bp Hind III fragment and cloned into a SP65 plasmid (10). RNA was

* To whom correspondence should be addressed

labelled by incorporation of ^{32}P-UTP into run-off transcripts using SP6 polymerase under conditions as described by the manufacturer. After cloning and sequencing of a human 5S rRNA gene in a 638 bp fragment (16) this gene was used for the synthesis of labelled RNA probes. This latter probe was used for the Southern analysis and for estimation of the 5S rRNA gene copy number.

Southern analysis

After digestion with restriction endonucleases, the DNA was electrophoresed in a 0.7% agarose gel and transferred to Hybond-N+ nylon membrane using 0.5 M NaOH, 0.5 M NaCl, followed by washing in 2×SSC. Hybridization was carried out in 0.5 M Na-phosphate (pH 7.2), 7% SDS, 1.0 mM EDTA at 65°C for 20 h. The membranes were washed twice at 22°C for 10 min in 0.04 M Na-phosphate (pH 7.2) 1% SDS and then digested with RNase A (5 μg/ml) for 30 min at 22°C. After further washing in 0.02 M Na-phosphate (pH 7.2) 0.5% SDS for 10 min at 22°C and 2 washings for 20 min in the same buffer at 75°C the membranes were slightly air dried and exposed to autoradiography at −70°C.

Subcloning and DNA sequencing

Subclones for sequencing were constructed in Bluescript M13 plasmids using appropriate restriction sites as shown in Fig. 3. One subclone, pHU5S3-360, used for in vitro transcription, was made by digestion of pHU5S3-638 with AccI and religation. This removed all 5-flanking sequences from the 5S gene since the first six nucleotides of the gene constitutes an AccI site. Sequencing was performed by the dideoxy method of Sanger (21) using ^{35}S-dATP and Sequenase or Taquence kits obtained from USB. Electrophoresis was performed in 6% polyacrylamide gels at 2500 V, the gels were dried in vacuum at 80°C for 1 h and autoradiography performed at −70°C.

Copy number

The total copy number of genes/gene variants and pseudogenes were determined in placenta DNA spotted onto nylon membranes using a dot spot apparatus. DNA was denatured in alkali transfer buffer (0.25 M NaOH, 1.5 M NaCl) and spotted in the concentration range 0.016−0.5 μg. Plasmid standards with or without a human 5S rRNA gene (pHU5S1 (Fig. 1) and Bluescript M13 plasmid DNA respectively) were denatured in the alkali transfer buffer and spotted in the range from 0.06 to 2.0 ng. To estimate the copy number in the 2.3 kb and 1.6 Kb fragments, 8 μg placenta DNA was digested with BamHI and electrophoresed in 0.7% agarose gels. Linearized plasmid pHU5S1 and Bluescript M13 plasmid DNA was loaded on the gel with 15 min intervals in varying amounts. The gel was blotted in alkali transfer buffer to Hybond-N$^+$ nylon membranes. Hybridization and washings were carried out as described above (Southern analysis) except for the last step which was performed at 65−70°C or at 75−80°C. The probe was prepared from the BamHI-SacI fragment from pHU5S1. After autoradiography the spots on the x-ray film were scanned and the hybridization to placenta DNA and plasmid DNA was compared. The copy number per haploid genome was estimated assuming the amount of DNA per cell to be 7 pg. DNA concentrations were measured by UV absorption, 2'-deoxyribose determination by the diphenylamine method (22) and comparisons of ethidium stained DNA in gels.

Transcription assay

The S100 extract was prepared from suspension cultures of HeLa cells as described by Dignam et al. (23). The protein content was 10 mg/ml. Transcription assays contained, in a final volume of 20 μl: 0.5 μg supercoiled plasmid DNA, 10 μl extract, 5 mM $MgCl_2$, 85 mM KCl, 10% glycerol, 10 mM Hepes, (pH 7.9), 0.5 mM ATP, GTP and CTP, 0.025 mM UTP, 10 μCi (α-^{32}P)UTP, 0.3 mM DTT, 0.3 mM PMSF (phenylmethylsulfonyl fluoride) and 0.1 mM EDTA. After incubation at 30°C for 60 min, RNA was phenol extracted and precipitated with EtOH from solutions containing 0.25 M NaCl, 25 mM Tris·base, pH 8.5, 25 mM EDTA, 0.5% SDS and 200 μg/ml tRNA. RNA was purified by three precipitations from 0.7 M NH_4OAc by addition of 3 vol EtOH. Samples for gel elctrophoresis were prepared by mixing 1 vol RNA in H_2O with 1 vol saturated urea and 1/3 vol XC-BB (xylene cyanol/bromophenol blue). Electrophoresis was performed in 10% polyacrylamide gels containing 7 M urea. Human 5S RNA was used as a size marker.

RESULTS

Cloning of 5S rRNA genes

Previous studies on the structure and organization of human 5S rRNA genes suggested that these are found in repeats of 2.8 kb (5). Recently a repeat lenght of 2.2 kb was found, and the sequence of one repeat was determined (15). We have cloned and partially characterized six restriction fragments containing human 5S rRNA genes. Human genomic DNA was digested with restriction endonucleases and electrophoresed in agarose gels. Restriction fragments in the 2.3 kb (SacI digests), 1.6 kb (BamHI

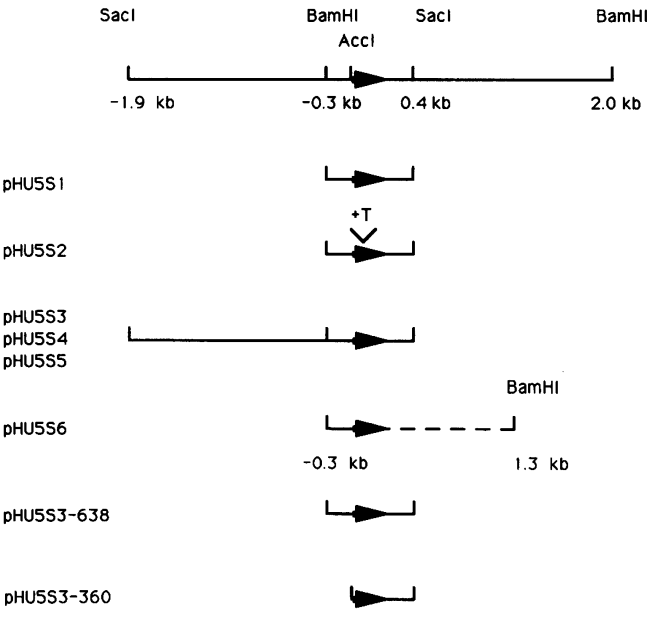

Figure 1. Diagram showing different cloned DNA fragments. These were obtained from human placenta DNA digested with BamHI+SacI (pHU5S1 and pHU5S2), SacI (pHU5S3, pHU5S4 and pHU5S5), or BamHI (pHU5S6). The BamHI-SacI fragments (638 bp), the SacI fragments (2.3 kb) and the BamHI fragment (1.6 kb) were cloned into Bluescript M13 plasmids. pHU5S3-638 and pHU5S3-360 were obtained by subcloning from pHU5S3. The 5S rRNA gene and gene variant are indicated by an arrow.

digests) and 640 bp (SacI + BamHI digests) regions were isolated and cloned into Bluescript M13 vectors. Six positive clones were isolated using a run-off transcript of a mouse 5S rRNA gene variant (10) as a probe . Two clones (pHU5S1 and pHU5S2) contained 640 bp BamHI-SacI inserts, three contained 2.3 kb SacI inserts (pHU5S3-pHU5S5) and one clone (pHU5S6) contained a 1.6 kb BamHI insert (Fig. 1).

Clones pHU5S1-pHU5S5 probably originate from a cluster of 2.3 kb tandem repeat units. Digestion of human genomic DNA with SacI (or BamHI) results in one major band of 2.3 kb hybridizing to a 5S rRNA probe as shown in Fig. 2, lane a. Limited digestion with this enzyme results in band sizes which are multiples of 2.3 kb, Fig. 2, lane b. These results are similar to what was found by others (15). After longer exposures of the films a smaller band of 1.6 kb appears after digestions of genomic DNA with BamHI but not with SacI (not shown). Clone pHU5S6 probably originates from this minor cluster.

Clones pHU5S1 and pHU5S3-pHU5S6 each contains a 5S rRNA gene whereas pHU5S2 contains a gene variant of 122 bp with an additional thymidine in position 47. The sequence of the BamHI-SacI fragment from pHU5S3 (pHU5S3-638, Fig.1) has been published (16). pHU5S6 contains a large deletion starting 12 bp 3′ of the gene, as compared to the 2.3 kb repeats.

The existence of any significant restriction lenght polymorphisms were ruled out by restriction analysis of DNA (lymphocyte) from 10 other persons (results not shown).

Copy number

The number of 5S rRNA genes/variant genes and pseudogenes were determined in total genomic DNA (Table 1). The denatured DNA was spottet onto nylon membrane using a dot spot apparatus and then hybridized to a ^{32}P-labelled run-off transcript from a human 5S rRNA gene (pHU5S1, Fig. 1). To distinguish between 5S genes, including closely related gene variants, and more divergent gene variants and pseudogenes, washings were carried out at two stringency intervals: low (65−70°C) and high (75−80°C). The total number of 5S related sequences per haploid genome is determined to be 1700−2000, which is in agreement with earlier findings (12). By raising the stringency this number is reduced to 300−400, demonstrating that the majority of 5S sequences in the human genome are pseudogenes or more divergent gene variants. We also determined the number of 5S sequences in the repeat structures. After digestion with restriction endonuclease BamHI the DNA was electrophoresed, blotted to nylon membranes and hybridized to the ^{32}P-labelled RNA probe. Washing at either low or high stringency resulted in a constant number of 100−150 hybridizing sequences in the 2.3 kb cluster (Table 1). Similarly, 5−10 sequences hybridized in the 1.6 kb cluster in the temperature interval from 65°C to 80°C. We conclude that these clusters contain only genes and gene

Table 1. Number of 5S rRNA genes/gene variants and pseudogenes in human genomic DNA

	Total genome	2.3 Kb repeat	1.6 Kb repeat
Low stringency of wash (65−70°C)			
Genes/gene variants + pseudogenes	1700−2000		
Genes/gene variants		100−150	5−10
High stringency of wash (75−80°C)			
Genes/gene variants	300−400	100−150	5−10

The total copy number of genes/gene variants and pseudogenes were determined per haploid genome. Placenta DNA was spotted onto nylon membranes using a dot spot apparatus. The number of genes/gene variants were meaured in the 2.3 Kb and 1.6 Kb repeats after digestion with BamHI, electrophoresis and blotting to nylon membranes. The low stringency wash (65−70°C) allow the RNA probe to remain hybridized to pseudogenes whereas only hybrids with genes and closely related gene variants are stable during washing at high stringency (75−80°C).

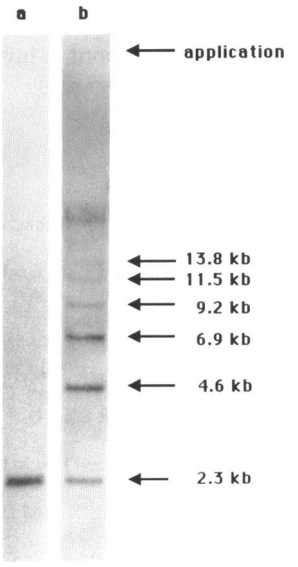

Figure 2. 5S rRNA genes in 2.3 kb repeats. Human genomic DNA from placenta was digested with restriction endonuclease SacI to completion (a) or to a limited extent (b).

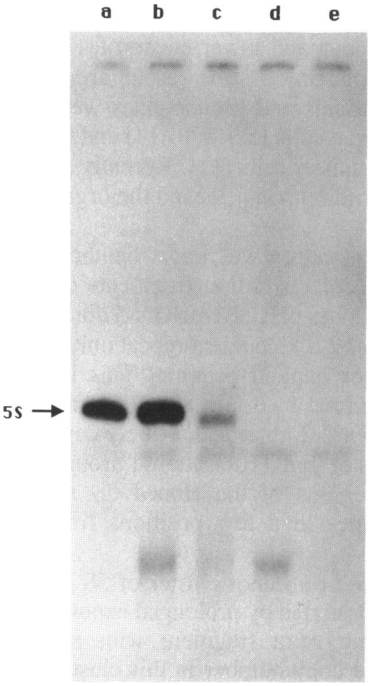

Figure 3. Transcription of 5S rRNA genes in a HeLa cell extract. 2.3 kb SacI fragment, pHU5S3 (a), 638 bp BamHI-SacI fragment, pHU5S3 (b), 360 bp AccI-SacI fragment, pHU5S3-360 (c), Bluescript M13 plasmid (d), no DNA (e). Assays a−c contained the same number of genes and were adjusted with bluescript plasmid DNA to the same total DNA concentration.

variants which can not be distinguished from the genes even at the high stringency washings. 200–300 genes/gene variants are not found in repeat structures but probably dispersed in the genome as single copies.

In vitro transcription

The transcription of human 5S genes has been studied by means of a synthetic gene without authentic flanking sequences (24,25). We have analysed the transcription of the isolated 5S genes and the gene variant in a HeLa S100 extract. The six clones supported the synthesis of 5S rRNA with similar efficiency. The results obtained by transcription of one of the clones, pHU5S3, containing the entire 2.3 kb repeat unit is shown in Fig. 3 (lane a). The transcription from clone pHU5S3 and subclone pHU5S3-638 (lane b), containing 273 bp 5′-flanking sequence, (Fig. 1) is of similar efficiency, whereas transcription from subclone pHU5S3-360 (lane c) without the 5′-flanking region, is reduced to less than 10% of the level of transcription from a gene with flanking sequences. This observation strongly indicates that sequences in the removed 5′-flanking region (up to position −273) are required for an efficient transcription. As controls we analysed the transcription of Bluescript M13 plasmid (lane d) and no DNA (lane e). The gene variant in pHU5S2, containing an extra thymidine in position 47, is transcribed as efficiently as the genes (not shown). The human 5S gene ICR has been identified through DNaseI protection studies with human TFIIIA (24,25). The additional thymidine is found within this region but apparently this position is not critical to the formation of a functional transcription complex.

DISCUSSION

The structural organization of 5S rRNA genes has been studied mainly in Xenopus and lower eucaryotes (Reviewed in 3,26) but from mammalian cells the results are scarce. Different 5S rRNA gene variants and pseudogenes were isolated from rat, mouse and human cells (5,7,8,10,14) and a bona-fide gene was isolated from hamster cells (11). Recently 5S rRNA genes were also isolated from human cells and the organization was studied (15).

In the present paper we report on the cloning and partial characterization of restriction fragments containing human 5S rRNA genes. Clones pHU5S1-pHU5S5 are suggested to originate from a cluster of 2.3 kb tandem repeat units with a copy number of 100–150 per haploid genome. This is probably the same cluster as described by Little and Braaten (15) who found 5S rRNA genes in a 2.2 kb repeat. Ninety copies of the gene were found in a cluster and 110 scattered around in the genome. In the flanking regions of the cloned 5S rRNA gene (15) the sequence deviates in a few positions from the pHU5S3-638 sequence (16).

The existence of a minor cluster of 5S genes in 1.6 kb repat units was demonstrated by prolonged exposures of southern blots and the cloning of a fragment with a corresponding size (pHU5S6). The copy number in this cluster was determined to be 5–10 per haploid genome. Like the major cluster the 1.6 kb cluster may contain both genes and closely related gene variants. In addition to the gene/gene variants we have found a large number of more distantly related gene variants and pseudogenes which brings the total number of 5S rRNA genes/gene variants and pseudogenes to 1700–2000 per haploid genome. The main cluster of 5S rRNA genes has been localized on chromosome nr. 1 in band q42.11-13 (27).

The role of flanking sequences in the transcription varies greatly for different 5S genes. In some studies it was found that the entire 5′-flanking region could be deleted from a Xenopus borealis somatic 5S gene without affecting the transcription in vitro (28). Others detected a roughly 2-fold reduction of transcription efficiency after substitution of nucleotides −11 to −16 upstream of this gene (29). An absolute requirement for 5′-flanking sequences has been reported for the transcription of D.melanogaster, B. mori and N.crassa 5S rRNA genes (30,31,32). A human synthetic gene without flanking sequences was efficiently transcribed in a homologous cell extract demonstrating that the coding region is sufficient for a certain level of transcription (24,25). We have found that removal of all 5′-flanking regions from an isolated human 5S gene results in a significant reduction in the level of transcription in a HeLa S100 extract. The exact sequences responsible for this effect are yet to be characterized, but two regions have drawn our attention in this context. One region with the sequence 5′-CGGGCGGG-GC-3′ (16) is located from position −35 to −44 and bears great resemblence to the GC-motif, a well known RNA polymerase II promoter element which binds transcription factor Sp1 (33). The GC-motif has a conserved hexanucleotide core sequence but a larger consensus sequence has been suggested (G/T)GGGCG-G(G/A)(G/A)(C/T) (Rewieved in 33). An Sp1-like binding site has previously been found 5′ to the RNA polymerase III transcribed EBER RNA (Epstein-Barr virus encoded small RNA) genes and shown to be involved in regulation of transcription (34). In addition we find that the upstream region of a hamster 5S rRNA gene (11) contains a GC motif (position −76 to −85), matching the consensus. Compared with the region following the human GC-like motif, the homology extends further downstream, resulting in a total of 15 conserved nucleotides.

The second region of interest is located around position −16 (16). In Xenopus 5S rRNA genes a strictly conserved pentamer-motif, AAAGT, at position −14 to −18, is suggested to serve as a binding signal for an upstream stimulatory factor (35,36). In the corresponding position upstream of the human 5S rRNA gene, the AG of this pentamer is conserved. It will be interesting to learn if this region or the GC motif is involved in regulation of transcription of the human 5S genes.

REFERENCES

1. Mao, J., Appel, B., Schaack, J., Sharp, S., Yamada, H. and Söll, D. (1982) Nucl. Acids Res. 10, 487–500.
2. Artavanis-Tsakonas, S., Schedl, P., Tschudi, C., Pirotta, V., Steward, R. and Gehring, W.J. (1977) Cell 12, 1057–1067.
3. Korn, L.J. (1982) Nature 295, 101–105.
4. Selker, E.U., Yanofsky, C., Driftmier, K., Metzenberg, R.L., Alzner-DeWeerd, B. and RajBhandary, U.L. (1981) Cell 24, 819–828.
5. Emerson, B.M. and Roeder, R.G. (1984) J. Biol. Chem. 259, 7916–7925.
6. Sharp, S., Garcia, A., Cooley, L. and Söll, D. (1984) Nucl. Acids Res. 12, 7617–7631.
7. Reddy, R., Henning, D., Rothblum, L. and Busch, H. (1986) J. Biol. Chem. 261, 10618–10623.
8. Arnold, G.J., Kahnt, B., Herrenknecht, K. and Gross, H.J. (1987) Gene 60, 137–144.
9. Lazar, E., Haendler, B. and Jacob, M. (1983) Nucl. Acids Res. 11, 7735–7741.
10. Leah, R., Frederiksen, S., Engberg, J. and Sørensen, P.D. (1990) Nucl. Acids Res. 18, 7441.
11. Hart, R.P. and Folk, W.R. (1982) J. Biol. Chem. 257, 11706–11711.
12. Hatlen, L. and Attardi, G. (1971) J. Biol. Chem. 56, 535–553.

13. Szabo, P., Lee, M.R., Elder, F.B. and Prensky, W. (1978) Chromosoma 65, 161–172.
14. Doran, J.L., Bingle. W.H. and Roy, K.L. (1987) Nucl. Acids Res. 15, 6297.
15. Little, R.D. and Braaten, D.C. (1989) Genomics 4, 376–383.
16. Sørensen, P.D., Simonsen, H. and Frederiksen, S. (1990) Nucl. Acids Res. 18, 3060.
17. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A laboratory manual, Cold Spring Harbor Laboratory, New York.
18. Holmes, D.S. and Quigley, M. (1981) Anal. Biochem. 114, 193–197.
19. Zasloff, M., Ginder, G.D. and Felsenfeld, G. (1978) Nucl. Acids Res. 5, 1139–1151.
20. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A laboratory manual, Cold Spring Harbor Laboratory, New York.
21. Sanger, F., Micklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
22. Burton, K. (1956) Biochem. J. 62, 315–323.
23. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucl. Acids Res. 11, 1475–1489.
24. Wingender, E., Frank, R., Blöcker, H., Wang, L., Jahn, D. and Seifart, K.H. (1988) Gene 64, 77–85.
25. Seifart, K.H., Wang, L., Waldschmidt, R., Jahn, D. and Wingender, E. (1989) J. Biol. Chem. 264, 1702–1709.
26. Geiduschek, E.P. and Tocchini-Valentini, G.P. (1988) Ann. Rev. Biochem. 57, 873–914.
27. Sørensen, P.D., Lomholt, B., Frederiksen, S. and Tommerup, N. (1990) Cytogenet. Cell Genet., in press.
28. Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980) Cell 19, 13–25.
29. Wormingon, W.M., Bogenhagen, D.F. and Brown, D.D. (1981) Cell 24, 809–817.
30. Sharp, S.J. and Garcia, A.D. (1988) Mol. Cell. Biol. 8, 1266–1274.
31. Morton, D.G. and Sprague, K.U. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 5519–5522.
32. Tyler, B.M. (1987) J. Mol. Biol. 196, 801–811.
33. Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) TIBS 11, 20–23.
34. Howe, J.G. and Shu, M. (1989) Cell, 57, 825–834.
35. Nietfeld, W., Digweed, M., Mentzel, H., Meyerhof, W., Köster, M., Knöchel, W., Erdmann, V.A. and Pieler, T. (1988) Nucl. Acids Res. 16, 8803–8815.
36. Oei, S. and Pieler, T. (1990) J. Biol. Chem. 265, 7485–7491.